Molecular Plant Breeding 2015, Vol.6, No.15, 1−22
http://mpb.biopublisher.ca

Research Article **Open Access**

Genetic Diversity Analysis in Tropical Maize Germplasm for Stem Borer and Storage Pest Resistance using Molecular Markers and Phenotypic traits

Mwololo J.K.[1], Munyiri S.W.[2], Semagn K.[3], Mugo S. [3], Okori p.[4]
1. Pwani University, P.O Box 195 Kilifi, Kenya
2. Chuka University, P.O. Box 109-60400, Chuka, Kenya.
3. International Maize and Wheat Improvement Center (CIMMYT), P. O. Box 1041, 00621, Nairobi, Kenya
4. International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), P. O. Box 1096, Lilongwe, Malawi, Kenya
Corresponding authors email: mwololojames@yahoo.com
Molecular Plant Breeding, 2015, Vol.6, No.15 doi: 10.5376/mpb.2015.06.0015
Received: 07 Jul., 2015
Accepted: 11 Aug., 2015
Published: 30 Sep., 2015

Preferred citation for this article:
Mwololo J.K., Munyiri S.W., Semagn K., Mugo S. and Okori p., 2015, Genetic Diversity Analysis in Tropical Maize Germplasm for Stem Borer and Storage Pest Resistance using Molecular Markers and Phenotypic traits, Molecular Plant Breeding, 6(15) 1−22(doi: 10.5376/mpb.2015.06.0015)

Abstract One hundred maize inbred lines and eighty four hybrids were characterized for resistance to maize stem borer and post-harvest insect pests. This was achieved using genetic distance and population structure based on simple sequence repeat (SSR) markers and biophysical traits. The test materials were evaluated for stem borer, maize weevil and larger grain borer (LGB) resistance. Leaf samples were harvested from 10 healthy plants per genotype and bulked. Genomic DNA was extracted using a modified version of mini-prep Cetyl Trimethyl Ammonium Bromide (CTAB) method. The samples were genotyped with 55 SSRs makers. Univariate analysis of variance was done using the general linear model procedure of SAS statistical package. Rodgers genetic distance was calculated for all data sets as a measure of genetic distance using NTSYS-pc for Windows. The distance matrices were used to generate phenograms using the unweighted pair group method based on arithmetic average (UPGMA) method in MEGA5. The genotypes were assigned into different populations using population structure software. The data was further subjected to discriminant and principal component analysis to group the gnotyoes. Analysis of molecular variance within and among the different populations was done using arlequin. There were significant differences ($P \leq 0.001$) for all the biophysical traits evaluated. The SSR marker data estimated successfully the close relationship among different hybrids and inbred lines within clusters. Comparisons of the different multivariate analyses revealed high concordance among the different approaches of analyses. The results of this study can be directly used by breeding programs to develop resistant genotypes.

Keywords Resistance; Maize insect pests; Genetic distance; Breeding, Molecular markerS

Introduction

Maize is a staple food for more than 300 million people in sub Saharan Africa (SSA) and is commonly grown by small-scale and resource poor farmers in rural areas (Shiferaw et al., 2011). However, the average maize yield in SSA was estimated at 1.4 t/ha, which is extremely low as compared to the 3.3 t/ha reported in developing countries in other parts of the world, the 4.9 t/ha worldwide production and the 8.4 t/ha in industrialized countries. Several factors, including a wide range of pests and diseases, periodic drought, scarcity of irrigation water, low soil fertility and farmers inability to use farm inputs contribute to low productivity in SSA. Insect pest in the field and in storage are among the factors that reduce yields and food availability in the region. Maize stem borers cause maize losses of up to 15% in susceptible germplasm in the infested ecologies, while the storage pest, such as maize weevil and larger grain borer (LGB) cause 20-30% yield loss (*http://www.syngentafoundation.org*). Although there are different possible methods that help in minimizing yield loss by insect pest (e.g. chemical, biological and cultural methods), host plant resistance developed through breeding is a preferred method to disseminate improved maize varieties due to its environmental and human safety, relatively low cost, and ease of use by farmers. However, there is very little effort in breeding for insect pest resistance in SSA which may be due to the genetic and logistical challenges associated with insect pest and hosts (screening and selecting for insect resistance). Nevertheless, CIMMYT and partners have developed various multiple borer resistance (MBR) lines and population using conventional breeding methods under artificial infestation. Some of the MBR germplasm have been released and disseminated in some countries.

Assessment of genetic relationship and population structure is an important tool that underpins successful breeding programs (Mohammadi and Prasanna, 2003; Mukhtar et al., 2002). Genetic distance is a measure of genetic divergence between species or between populations within a species. Smaller genetic distances indicate a close genetic relationship whereas large genetic distances indicate a more distant genetic relationship. In a breeding program, genetic gain achieved through artificial selection is proportional to the extent of genetic differences present in the parental lines or populations. Thus, the correct choice of parents can influence the outcome of selection (Bohn et al., 1999). Depending on the objectives of a breeding programme, breeders use different methods in selecting the best parental combinations, including (a) pedigree relationships, (b) morphological and agronomic traits, (c) adaptability and yield stability, and (d) genetic distances estimated from morphological and molecular markers (Bohn et al. 1999; Maric et al., 2004; Bertan et al., 2007). Morphological and agronomic traits were the earliest genetic markers used in germplasm characterization and quantifying genetic distance in crops but they have a number of limitations including low polymorphism, low heritability, late expression during the development process and are highly influenced by the environment (Smith and Smith 1989).

In contrast, molecular markers, are more effective than morphological and agronomic traits for germplasm characterization. Genetic distance and population structure can be estimated from various types of molecular markers, including restriction fragment length polymorphism (RFLP), amplified fragment length polymorphism (AFLP), random polymorphic DNA (RAPD), microsatellites or simple sequence repeats (SSRs) and single nucleotide polymorphisms (SNPs). SSR makers are widely used by maize researchers because they are available in large numbers in the public domain (*MaizeGDB: http://www.maizegdb.org)*, co-dominant, multiallelic, highly polymorphic even in closely related individuals, can be exchanged between laboratories, and have uniform distribution in the genome (Gupta et al., 2002; Prasanna et al., 2010). Although advances in marker technology have shifted toward SNP markers, particularly for model organisms with substantial genomic resources, SSRs markers perform better at clustering germplasm into populations and providing more resolution in measuring genetic distance than SNPs markers (Hamblin et al., 2007).

Genetic variability for resistance to field and postharvest insect pests using phenotypic data have been reported (Munyiri et al., 2010; Tefera, 2012). However, the extent of genetic differences and patterns of relationships among this germplasm and its response to stem borer, weevil and LGB resistance has not been well studied. The objective of this study was therefore to understand the extent of genetic difference, relationship and population structure across a subset of tropical maize germplasm that has been bred for field and storage pests' resistance using SSR markers and biophysical traits.

1 Results

1.1 Phenotypic evaluation

There were significant differences ($P \leq 0.001$) among the maize inbred lines and hybrids for all the biophysical and bioassay traits measured in the study. These traits were used to group the maize germplasm into resistant and susceptible.

1.2 Genetic distance and relationship

Roger's genetic distance between pairwise comparisons of all the 184 genotypes ranged from 0.004 to 0.467, and the overall average distance was 0.302. The vast majority (92.4 %) fell between 0.200 and 0.400 (Figure 1).

The UPGMA tree generated from Roger's genetic distance matrix grouped the majority of the genotypes into two major groups, one for inbred lines and the other for hybrids (Figure 2). The first group had three sub-groups (NA, G1 and G2) while the second group had also three sub-groups G3, G4 AND G5. Sub-group one (G1) consisted of a total of 68 inbred lines, including

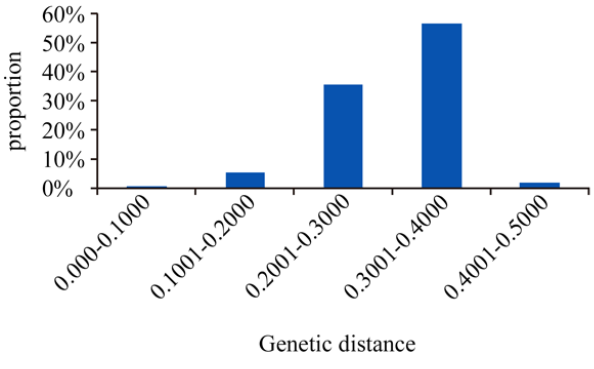

Figure 1 Frequency histogram of the different genotypes based on genetic distance

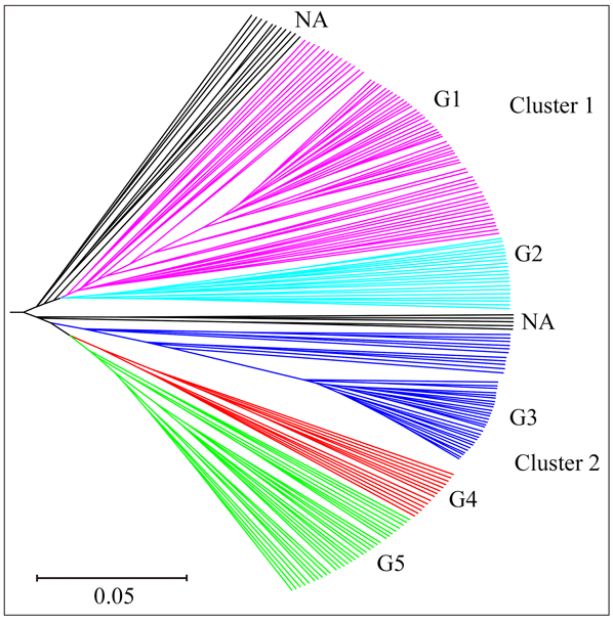

Figure 2 UPGMA tree for 100 inbred lines and 84 hybrids based on Roger's genetic distance calculated from 227 SSR alleles. The different groups are indicated with different colours, and detail group membership is provided in appendix 1

29 lines that are resistant both for storage pests and stem borers, 23 lines that are resistant only to storage pests, 14 lines that are resistant only to stem borers, and 2 lines that are susceptible to both storage and field pests.

Sub-group 2 (G2) consisted of inbred lines which have been bred for both stem borer and storage insect pests (9 lines), stem borer resistance (15 lines) and yield (2 lines).

In the second group which was composed of hybrids, Sub-group 3 (G3) consisted of hybrids which had been bred for storage pest resistance (23 hybrids), stem borers (10 hybrids) and grain yield (5 hybrids).

Group 4 (G4) was composed of 13 commercial hybrids from different seed companies which were all susceptible to the storage insect, and only 4 of the 13 hybrids showed some levels of resistance to the stem borer.

Group five (G5) consisted of 25 hybrids that were resistant to stem borer and two hybrids resistant to both stem borer and the storage insect pests.

The first five principal components from principal component analysis explained 25.7% of the total SSR variations among samples. A plot of PC1 (8.8%) and PC2 (7.4%) revealed 3 major groups (Figure 3) and the pattern of grouping was the same as for the model-based population partition at k=3.

1.3 The population structure based on SSRs

The estimated log probability of the data (LnP(D)) increased sharply between K = 1 and K = 4 (Figure 4b), and fairly stabilized between K = 5 and K = 6 (Figure 4a). The ad hoc statistic ΔK showed a higher likelihood values at K = 3 (Figure 4b), with a sharp decrease when K increased from 3 to 6 (Figure 4a). Therefore the estimated LnP (D) and K both suggest the presence of 3 possible groups.

Assignment of genotypes into specific groups was irrespective of the type of germplasm (inbred versus hybrids) and generally followed their pedigree information and their reaction to field and storage pests, with overlapping variation with some other traits, such as grain yield and drought tolerance. The majority of the genotypes were assigned to group 2, which included 23 hybrids (CKIR series) and 15 inbred lines (CKSB series) bred for stem borer resistance, 18 commercial hybrids and other inbred lines from the CIMMYT breeding programs. Group 1 and 3 consisted of 41 inbred lines in CKSP series and 28 hybrids in CKPH series that were bred for storage pest resistance within the CIMMYT breeding program. The mixed population generally was made up of CIMMYT inbred lines bred for yield and drought tolerance.

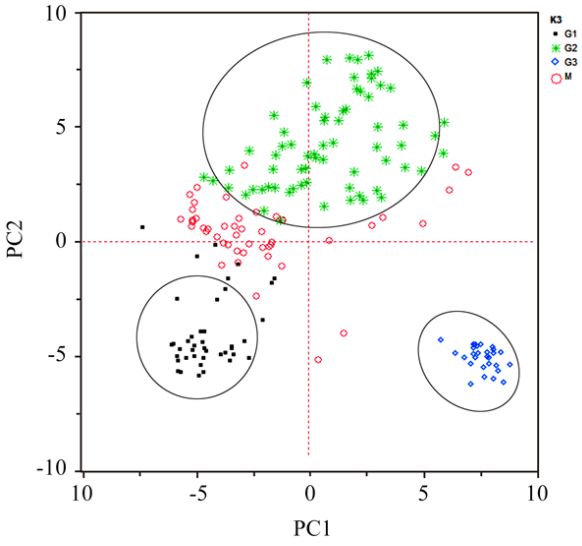

Figure 3 Principal component analysis (PCA) of 184 genotypes based on 56 SSRs. The groups from PCA supports the presence of population structure at K=3. Individuals that were assigned in to a mixed group in the population structure analysis are indicated in circle (red colour)

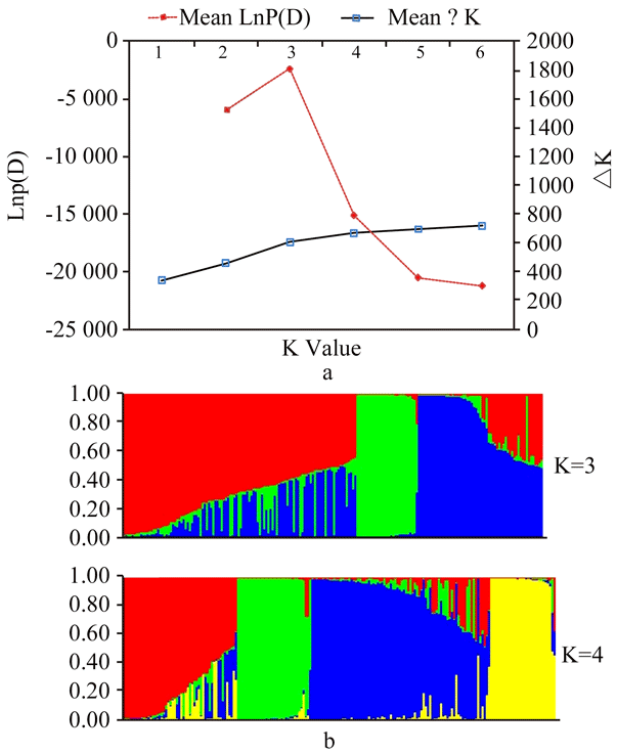

Figure 4 Population structure of 184 genotypes based on 227 alleles from 56 SSR markers: a) Plot of LnP (D) and ΔK calculated for K ranging between 1 and 6, with each K represented by a mean of 3 repeats

b) Population structure of the 184 genotypes at K =3 and K =4. Each individual is represented by a single vertical line that is partitioned into K coloured segments, with lengths proportional to the estimated probability membership to each of the K inferred clusters

1.4 Discriminant analysis

The reliability of the different groups obtained through the model-based population structure and cluster analyses was assessed through discriminant analyses using the group membership from both methods as categorical variables. The discrimination model with the stepwise procedure identified 12 alleles from 11 SSRs as the best explanatory variables for the priori group defined at K = 3 and 22 alleles from 21 SSRs for the prior groups obtained using cluster analysis (Table 1 and 2 shows the list of SSR alleles that were chosen by the stepwise discriminant analyses). The Mahalanobis distance matrix from pairwise comparisons of the 3 groups obtained from STRUCTURE at K=3 ranged from 4.0 to 37.0 and they were all significant, with group 3 being 2 to 11 times more distant from all others.

The Mahalanobis distance between groups obtained using cluster analysis ranged from 9.84 to 83.4. The commercial hybrids (CHS) were generally more distant from all the other genotypes. Based on the population structure, the grouping at K=3 corresponds to the clustering based on the Rodgers genetic distance since population 1 was equivalent to the SPRL, population two constituted the SBRL and SBRH which were close to one another with a distance of 9.84 between them, and the commercial hybrids (G4 in the dendrogram), population 3 to SPRH and the mixed population constituted other CIMMYT lines bred for yield and drought. The phenotypic traits for classifying the genotypes into resistance and susceptible was not a good indicator for discriminating the genotypes, since the canonical correlation coefficient (CAN1) was 0.13 and 0.26 for the stem borer and storage pest resistance indices respectively.

Comparisons of the different multivariate analyses revealed high concordance among the PCA, model-based population partition, clustering based on the genetic distance and discriminant analyses in terms of the number of groups and members in each group. However, there was low concordance between grouping based on the phenotypic data indices and the SSR based population partitioning in assigning the genotypes into the different groups or populations.

1.5 Analysis of molecular variance (AMOVA)

Table 3 shows the partitioning of the overall SSR variance into hierarchical levels using AMOVA. When AMOVA was performed using the 6 possible groups predicted from UPGMA-cluster analyses and population structure; and the two groups based on storage pest resistance, the estimated fixation indices (FST) varied from 6.49 % to 27.85%. When the overall SSR variance was partitioned into hierarchical levels using the groups predefined from the model-based population partition at K = 2, K = 3, K = 4, K = 5 and K = 6 as categorical variables, FST accounted for 15.3%, 23.8%, 25.86%, 26.56% and 27.85%, respectively. In the cluster analysis that based on the storage pest resistance trait, FST accounted for 24.26% and 6.49% respectively. A random permutation test indicated that the proportion of variances attributable at all groups were highly significant ($p < 0.0001$).

Table 1 List of SSR alleles that were chosen by the stepwise discriminant analyses for groups at K=3

No. of Variable	Variables	Variable IN/OUT	Status	Patial R2	F	Pr > F	Wilks Lambda	Pr < Lambda
1	umc1061_109	umc1061_109	IN				0.3850	<0.0001
2	phi029_147 /…	phi029_147	IN	0.524	65.76	<0.0001	0.1830	<0.0001
3	phi008_54 /…	phi008_54	IN	0.358	33.11	<0.0001	0.1180	<0.0001
4	phi008_54/…	phi109275-138	IN	0.270	21.84	<0.0001	0.0860	<0.0001
5	phi008_54/..	phi453121_217	IN	0.247	19.28	<0.0001	0.0650	<0.0001
6	phi008_54/…	phi059_157	IN	0.228	17.18	<0.0001	0.0500	<0.0001
7	phi008_54/…	umc1136_136	IN	0.192	13.75	<0.0001	0.0400	<0.0001
8	phi008_54/..	umc1143_83	IN	0.163	11.20	<0.0001	0.0340	<0.0001
9	phi008_54/…	phi109188_167	IN	0.215	15.73	<0.0001	0.0270	<0.0001
10	phi008_54/	phi114_166	IN	0.152	10.23	<0.0001	0.0220	<0.0001
11	phi008_54/..	phi051_139	IN	0.135	8.87	<0.0001	0.0190	<0.0001
12	phi008_54/…	umc1061_112	IN	0.119	7.58	<0.0001	0.0170	<0.0001
13	phi008_54/…	phi079_178	IN	0.110	6.92	0.000	0.0150	<0.0001
14	phi008_54/..	phi072_153	IN	0.110	6.85	0.000	0.0140	<0.0001
15	phi008_54/	umc1143_78	IN	0.093	5.65	0.001	0.0120	<0.0001
16	phi008_54/	phi029_151	IN	0.102	6.22	0.000	0.0110	<0.0001
17	phi008_54/..	phi051_136	IN	0.096	5.77	0.001	0.0100	<0.0001
18	phi008_54/	phi127_113	IN	0.089	5.34	0.002	0.0900	<0.0001
19	phi008_54/	phi014_159	IN	0.102	6.11	0.001	0.0800	<0.0001
20	phi008_54/..	phi96100_297	IN	0.094	5.59	0.001	0.0700	<0.0001
21	phi008_54/	phi059_154	IN	0.092	5.37	0.002	0.0700	<0.0001
22	phi008_54/	phi112_133	IN	0.090	5.22	0.002	0.0600	<0.0001
23	phi008_54/..	phi114_134	IN	0.085	4.92	0.003	0.0600	<0.0001
24	phi008_54/…	phi064_71	IN	0.105	6.14	0.001	0.0500	<0.0001
25	phi008_54/…	umc2047_126	IN	0.083	4.73	0.003	0.0500	<0.0001

Continuing table 1

No. of Variable	Variables	Variable IN/OUT	Status	Patial R2	F	Pr > F	Wilks Lambda	Pr < Lambda
26	phi008_54/..	phi014_162	IN	0.072	4.04	0.008	0.0400	<0.0001
27	phi008_54/…	phi089_87	OUT	0.084	4.74	0.003	0.0400	<0.0001
28	phi008_54/…	phi075_238	IN	0.074	4.06	0.008	0.0400	<0.0001
29	phi008_54/..	umc1136_136	IN	0.018	0.94	0.424	0.0400	<0.0001
30	phi008_54/…	phi015_95	IN	0.068	3.70	0.013	0.0300	<0.0001
31	phi008_54/…	phi112_149	IN	0.067	3.64	0.014	0.0300	<0.0001
32	phi008_54/..	phi115_301	IN	0.064	3.45	0.018	0.0300	<0.0001
33	phi008_54/…	phi062_164	IN	0.073	3.96	0.009	0.0300	<0.0001
34	phi008_54/…	phi102228_128	IN	0.082	4.41	0.005	0.0300	<0.0001
35	phi008_54/..	phi041_191	OUT	0.060	3.14	0.027	0.0200	<0.0001
36	phi008_54/…	phi084_151	IN	0.066	3.46	0.018	0.0200	<0.0001
37	phi008_54/…	umc2047_130	IN	0.078	4.14	0.008	0.0200	<0.0001
38	phi008_54/..	umc1161_137	IN	0.063	3.22	0.025	0.0200	<0.0001
39	phi008_54/…	phi051_139	IN	0.040	2.03	0.112	0.0200	<0.0001
40	phi008_54/…	phi031_189	IN	0.061	3.12	0.028	0.0200	<0.0001
41	phi008_54/..	phi064_87	IN	0.076	3.95	0.010	0.0200	<0.0001
42	phi008_54/…	phi109188_155	IN	0.075	3.84	0.011	0.0200	<0.0001
43	phi008_54/…	phi084_154	IN	0.073	3.70	0.013	0.0200	<0.0001
44	phi008_54/..	phi053_177	IN	0.073	3.71	0.013	0.0100	<0.0001
45	phi008_54/…	phi041_195	IN	0.088	4.48	0.005	0.0100	<0.0001
46	phi008_54/…	phi075_228	IN	0.065	3.20	0.025	0.0100	<0.0001
47	phi008_54/..	phi053_169	IN	0.065	3.19	0.026	0.0100	<0.0001
48	phi008_54/…	phi084_160	IN	0.066	3.25	0.024	0.0100	<0.0001
49	phi008_54/…	phi453121_223	IN	0.068	3.32	0.022	0.0100	<0.0001
50	phi008_54/..	umc1917_141	OUT	0.078	3.82	0.011	0.0100	<0.0001
51	phi008_54/…	phi308707_116	IN	0.073	3.53	0.017	0.0100	<0.0001

Continuing table 1

No. of Variable	Variables	Variable IN/OUT	Status	Patial R2	F	Pr > F	Wilks Lambda	Pr < Lambda
52	phi008_54/…	umc1917_132	IN	0.064	3.02	0.032	0.0100	<0.0001
53	phi008_54/	phi109188_163	IN	0.064	3.00	0.033	0.0100	<0.0001
54	phi008_54/	phi104127_169	IN	0.060	2.77	0.044	0.0100	<0.0001
55	phi008_54/	phi056_239	IN	0.059	2.74	0.046	0.0100	<0.0001
56	phi008_54/	phi014_159	IN	0.044	1.99	0.119	0.0100	<0.0001
57	phi008_54/	umc1332_143	IN	0.065	2.99	0.033	0.0100	<0.0001
58	phi008_54/	phi085_240	IN	0.072	3.33	0.022	0.0100	<0.0001
59	phi008_54/	umc1196_135	IN	0.071	3.25	0.024	0.0000	<0.0001
60	phi008_54/	phi96100_273	IN	0.071	3.23	0.025	0.0000	<0.0001
61	phi008_54/	umc1061_106	IN	0.074	3.36	0.021	0.0000	<0.0001
62	phi008_54/	umc1136_157	IN	0.067	2.97	0.035	0.0000	<0.0001
63	phi008_54/	phi085_255	IN	0.073	3.25	0.024	0.0000	<0.0001
64	phi008_54/	phi072_161	IN	0.078	3.46	0.019	0.0000	<0.0001
65	phi008_54/	umc2047_114	IN	0.079	3.51	0.017	0.0000	<0.0001
66	phi008_54/	umc1136_151	IN	0.072	3.13	0.028	0.0000	<0.0001
67	phi008_54/	phi014_156	IN	0.067	2.85	0.040	0.0000	<0.0001
68	phi008_54/	umc1136_139	IN	0.069	2.96	0.035	0.0000	<0.0001

* The chosen alleles had a P value of Pr < F 0.0001

Table 2 List of SSR alleles that were chosen by the stepwise discriminant analyses for groups from cluster analysis

No.Of Variable	Variable	Variable IN/OUT	Status	Patial R^2	F	Pr > F	Wilks Lambda	Pr < Lambda
1	phi114_158	phi114_158	IN				0.2600	<0.0001
2	phi114_158/..	umc1545_84	IN	0.440	27.88	<0.0001	0.1500	<0.0001
3	phi114_158/..	umc1061_100	IN	0.380	21.26	<0.0001	0.0900	<0.0001
4	phi114_158/..	phi308707_116	IN	0.340	18.16	<0.0001	0.0600	<0.0001
5	phi114_158/..	umc1447_113	IN	0.300	15.04	<0.0001	0.0400	<0.0001
6	phi114_158/..	umc1196_141	IN	0.270	12.90	<0.0001	0.0300	<0.0001
7	phi114_158/..	umc1136_136	IN	0.250	11.53	<0.0001	0.0200	<0.0001
8	phi114_158/..	phi014_162	IN	0.230	10.37	<0.0001	0.0200	<0.0001
9	phi114_158/..	umc2047_126	IN	0.230	10.01	<0.0001	0.0100	<0.0001
10	phi114_158/..	phi453121_205	IN	0.210	9.20	<0.0001	0.0100	<0.0001
11	phi114_158/..	phi96100_297	IN	0.270	12.30	<0.0001	0.0100	<0.0001
12	phi114_158/..	phi127_125	IN	0.220	9.16	<0.0001	0.0100	<0.0001
13	phi114_158/..	phi008_54	IN	0.180	7.25	<0.0001	0.0100	<0.0001
14	phi114_158/..	phi029_147	IN	0.170	6.80	<0.0001	0.0000	<0.0001
15	phi114_158/..	phi109275_130	IN	0.170	6.86	<0.0001	0.0000	<0.0001
16	phi114_158/..	phi059_157	IN	0.150	5.94	<0.0001	0.0000	<0.0001
17	phi114_158/..	phi96100_293	IN	10.170	6.62	<0.0001	0.0000	<0.0001
18	phi114_158/..	umc1061_1009	IN	0.170	6.49	<0.0001	0.0000	<0.0001
19	phi114_158/..	phi109188_171	IN	0.180	6.92	<0.0001	0.0000	<0.0001
20	phi114_158/..	umc1136_139	IN	0.160	5.87	<0.0001	0.0000	<0.0001
21	phi114_158/..	phi059_145	IN	0.150	5.64	<0.0001	0.0000	<0.0001
22	phi114_158/..	phi093_282	IN	0.160	6.01	<0.0001	0.0000	<0.0001
23	phi114_158/..	phi96100_265	IN	0.140	5.07	0.000	0.0000	<0.0001
24	phi114_158/..	phi031_217	IN	0.140	4.92	0.000	0.0000	<0.0001
25	nc133_108/p…	nc113_108	IN	0.140	4.85	0.000	0.0000	<0.0001
26	nc133_108/p…	umc1447_125	IN	0.150	5.32	0.000	0.0000	<0.0001
27	nc133_108/p…	umc1136_157	OUT	0.150	5.23	0.000	0.0000	<0.0001
28	nc133_108/p…	phi96100_273	IN	0.140	5.00	0.000	0.0000	<0.0001
29	nc133_108/p…	phi015_103	IN	0.140	4.98	0.000	0.0000	<0.0001
30	nc133_108/p…	phi064_71	IN	0.150	5.10	0.000	0.0000	<0.0001
31	nc133_108/p…	umc1136_133	IN	0.130	4.55	0.000	0.0000	<0.0001
32	nc133_108/p…	umc1143_83	IN	0.120	4.16	0.000	0.0000	<0.0001
33	nc133_108/p…	phi059_151	IN	0.120	4.04	0.000	0.0000	<0.0001
34	nc133_108/p…	phi059_154	IN	0.120	4.06	0.000	0.0000	<0.0001
35	nc133_108/p…	phi227562_323	OUT	0.120	3.97	0.000	0.0000	<0.0001
36	nc133_108/p…	phi063_174	IN	0.120	3.80	0.000	0.0000	<0.0001
37	nc133_108/p…	phi115_289	IN	0.120	3.92	0.000	0.0000	<0.0001
38	nc133_108/p…	umc2047_134	IN	0.130	4.05	0.000	0.0000	<0.0001
39	nc133_108/p…	phi056_248	IN	0.110	3.36	0.010	0.0000	<0.0001
40	nc133_108/p…	phi112_151	IN	0.110	3.47	0.010	0.0000	<0.0001
41	nc133_108/p…	phi104127_157	IN	0.110	3.54	0.000	0.0000	<0.0001
42	nc133_108/p…	phi123_146	IN	0.120	3.66	0.000	0.0000	<0.0001

Continuing table 2

No.Of Variable	Variable	Variable IN/OUT	Status	Patial R^2	F	Pr > F	Wilks Lambda	Pr < Lambda
43	nc133_108/p…	phi075_238	IN	0.110	3.50	0.010	0.0000	<0.0001
44	nc133_108/p…	umc1143_83	OUT	0.060	1.66	0.150	0.0000	<0.0001
45	nc133_108/p…	phi076_159	IN	0.120	3.71	0.000	0.0000	<0.0001
46	nc133_108/p…	phi014_162	OUT	0.050	1.45	0.210	0.0000	<0.0001
47	nc133_108/p…	phi022_134	IN	0.110	3.36	0.010	0.0000	<0.0001
48	nc133_108/p…	phi102228_128	IN	0.110	3.45	0.010	0.0000	<0.0001
49	nc133_108/p…	phi374118_226	IN	0.120	3.52	0.010	0.0000	<0.0001
50	nc133_108/p…	phi056_245	OUT	0.110	3.37	0.010	0.0000	<0.0001
51	nc133_108/p…	umc1061_103	IN	0.100	3.08	0.010	0.0000	<0.0001
52	nc133_108/p…	phi104127_157165	IN	0.120	3.61	0.000	0.0000	<0.0001
53	nc133_108/p…	phi041_195	IN	0.110	3.08	0.010	0.0000	<0.0001
54	nc133_108/p…	phi102228_132	IN	0.130	3.83	0.000	0.0000	<0.0001
55	nc133_108/p…	umc1136_136	OUT	0.060	1.68	0.140	0.0000	<0.0001
56	nc133_108/p…	phi063_154	IN	0.130	3.73	0.000	0.0000	<0.0001
57	nc133_108/p…	phi227562_317	IN	0.120	3.40	0.010	0.0000	<0.0001
58	nc133_108/p…	phi072_137	IN	0.120	3.54	0.000	0.0000	<0.0001
59	nc133_108/p…	umc1061_106	IN	0.110	3.24	0.000	0.0000	<0.0001
60	nc133_108/p…	phi051_142	IN	0.120	3.32	0.010	0.0000	<0.0001
61	nc133_108/p…	umc1196_153	IN	0.140	3.99	0.010	0.0000	<0.0001
62	nc133_108/p…	phi050_82	IN	0.100	2.83	0.000	0.0000	<0.0001
63	nc133_108/p…	phi064_103	IN	0.100	2.82	0.020	0.0000	<0.0001
64	nc133_108/p…	phi085_240	IN	0.100	2.79	0.020	0.0000	<0.0001
65	nc133_108/p…	phi014_159	IN	0.100	2.70	0.020	0.0000	<0.0001
66	nc133_108/p…	umc1136_154	IN	0.100	2.77	0.020	0.0000	<0.0001
67	nc133_108/p…	umc1136_136	IN	0.090	2.41	0.040	0.0000	<0.0001
68	nc133_108/p…	phi102228_124	IN	0.090	2.33	0.050	0.0000	<0.0001
69	nc133_108/p…	phi053_177	IN	0.090	2.42	0.040	0.0000	<0.0001
70	nc133_108/p…	phi108411_119	IN	0.090	2.31	0.050	0.0000	<0.0001
71	nc133_108/p…	phi084_157	IN	0.110	2.76	0.020	0.0000	<0.0001
72	nc133_108/p…	umc1917_132	IN	0.090	2.30	0.050	0.0000	<0.0001
73	nc133_108/p…	phi227562_323	OUT	0.080	1.89	0.010	0.0000	<0.0001
74	nc133_108/p…	phi084_151	IN	0.100	2.44	0.040	0.0000	<0.0001
75	nc133_108/p…	phi056_248	OUT	0.060	1.52	0.190	0.0000	<0.0001
76	nc133_108/p…	phi064_87	IN	0.100	2.52	0.030	0.0000	<0.0001
77	nc133_108/p…	phi085_240	OUT	0.080	1.87	0.100	0.0000	<0.0001

* The chosen alleles had a P value of Pr<F 0.0001

2 Discussion

The significant differences and wide range in the means of the phenotypic traits related to resistance among the germplasm shows that there is great potential for the development of improved maize genotypes that are resistant to the postharvest insect pests. The biophysical/bioassay and molecular data confirm the existence of genetic divergence in tropical maize germplasm in response to the maize field and storage insect pests. This is in agreement with earlier studies that reported the existence of genetic variability of resistance to the maize weevil, larger

Table 3 Analysis of molecular variance (AMOVA) for the extraction of SSR variation among groups (populations) and among individuals within populations

Grouping	Source of variation	d.f.	Sum of squares	Variance components	Percentage of variation
K=2 (2 pos and mixed)	Among populations	2	521.378	5.3344	15.3*
	Within populations	181	5326.046	29.42567	84.65
	Total	183	5847.424	34.76006	
K=3 (3 pops and mixed	Among populations	3	1156.175	8.07679	23.66*
	Within populations	180	4691.249	26.06249	76.34
	Total	183	5847.424	34.139	
K=4 (4 pops and mixed)	Among populations	4	1349.915	8.7638	25.86*
	Within populations	179	4497.509	25.12575	74.14
	Total	183	5847.424	33.88954	
K=5 (5 pops and mixed)	Among populations	5	1450.722	8.93096	26.56*
	Within populations	178	4396.702	24.70058	73.44
	Total	185	5847.424	33.63154	
K=5 (6 pops and mixed)	Among populations	6	1566.737	9.3348	27.85*
	Within populations	177	4280.687	24.18467	72.15
	Total	183	5847.424	33.51947	
Cluster analysis (6 groups)	Among populations	5	1303.463	8.16166	24.26*
	Within populations	177	4509.296	25.47625	75.74
	Total	182	5812.76	33.63791	
Analysis based on SPR(2 groups)	Among populations	1	226.525	2.14151	6.49*
	Within populations	181	5586.235	30.86318	93.51
	Total	182	5812.76	33.00469	

*p-value<0.0001

grain borer and the stem borers among tropical maize germplasm (Arnason et al. 1994; Mwololo et al., 2010; Tefera et al., 2011). This genetic diversity can be exploited in breeding programs to introgress resistance to field and postharvest insect pests into improved varieties using conventional and genetic engineering approaches (Dhliwayo and Pixley 2003).

Overall mean Roger's genetic distance of 0.353 among pairwise comparisons of inbred lines, with the vast majority (94.2 %) showing distances between 0.300 and 0.400 have been reported (Semagn et al., 2012). This slightly differs from the average distance (0.3012) obtained from the current study. The observed lower genetic distance is likely due to the mixed origin of the inbred lines and hybrids. Clustering of the individual candidates among the wide germplasm evaluated in relation to resistance to the maize stem borer and postharvest insect pests was evident. Some of the genotypes which had been bred for stem borer and storage insect pests were resistant to both classes of maize insects hence has the potential to breed for multiple resistance. In addition, the clustering based on the SSR marker conforms to the history of generating the different genotypes. The grouping based on the phenotypic traits did not show a clear genetic differentiation with regard to specific resistance traits of the six different groups from the cluster analysis based on the SSR marker data. This is in agreement with previous studies whereby there was lack of clear clustering patterns based on phenotypes, environmental adaptation and grain colour (Xia et al., 2005). This can be explained by the fact that, selectively neutral markers used were not subject to selection and thus resistance, an adaptive trait had low correlation with SSR data (Koebner et al., 2002). The

Appendix 1 Summary of SSR markers used in the genetic relationship study among the 184 maize genotypes

Locus/Marker	Bin position	Repeat length	Allele size range (bp)	No. of alleles for inbred lines (n=115)	No. of alleles for hybrids (n=83)	No. of alleles for OPVs and landraces (n=84)	Total number of alleles
nc133	2.05	5	106-113	2	2	2	2
phi008	5.03	3	55-96	8	5	1	9
phi014	8.04	3	147-162	3	4	4	4
phi015	8.09	4	82-108	5	7	7	7
phi022	9.03	4	134-166	2	2	1	2
phi029	3.04	2 and 4	144-158	3	5	4	7
phi031	6.04	4	185-222	5	5	5	5
phi041	10.00	4	187-210	5	4	5	6
phi050	10.03	4	73-84	3	1	1	4
phi051	7.05	3	135-144	3	3	2	4
phi053	3.05	4	170-195	3	4	3	4
phi056	1.01	3	234-244	5	5	5	7
phi059	10.02	3	124-158	5	5	6	6
phi062	10.04	3	160-164	2	1	1	2
phi063	10.02	4	150-222	5	4	8	9
phi064	1.11	4	72-107	7	9	9	9
phi072	4.01	4	136-161	4	7	6	7
phi075	6.00	2	222-239	5	5	4	8
phi076	4.11	6	153-176	3	5	5	5
phi079	4.05	5	179-193	4	3	3	4
phi084	10.04	3	148-160	4	4	4	5
phi085	5.07	5	230-257	4	4	5	6
phi089	6.08	4	86-94	2	2	2	2
phi090	2.08	5	138-146	1	1	1	2
phi093	4.08	4	281-292	3	3	2	4
phi102228	3.04	4	100-132	2	0	2	4
phi104127	3.01	4	152-169	3	3	4	5
phi108411	9.06	4	110-139	3	1	2	7
phi109188	5.00	4	154-170	4	3	4	5
phi109275	1.00	4	51-137	7	10	6	11
phi112	7.01	2	133-158	4	3	2	5
phi114	7.02	4	128-169	5	5	5	6
phi115	8.03	2 and 4	280-301	2	2	2	5
phi123	6.07	4	142-146	2	2	2	2
phi127	2.08	4	109-129	3	1	4	6
phi227562	1.12	3	304-327	6	4	7	8
phi308707	1.10	3	109-131	4	5	5	5
phi331888	5.04	3	129-135	3	3	3	3
phi374118	3.02	3	205-233	4	6	6	7
phi453121	3.01	3	206-224	3	4	4	4

Continuing Appendix 1

Locus/Marker	Bin position	Repeat length	Allele size range (bp)	No. of alleles for inbred lines (n=115)	No. of alleles for hybrids (n=83)	No. of alleles for OPVs and landraces (n=84)	Total number of alleles
phi96100	2.00	4	266-298	5	7	8	9
umc1061	10.06	3	100-112	2	4	3	5
umc1136	3.10	3	130-157	6	7	5	7
umc1143	6.00	5	73-85	3	3	3	3
umc1161	8.06	6	136-149	3	2	3	3
umc1196	10.07	6	133-158	4	4	4	5
umc1266	3.06	3	120-144	3	0	1	3
umc1304	8.02	4	124-133	2	1	5	3
umc1332	5.04	3	116-143	5	5	1	6
umc1367	10.03	3	146-159	0	0	3	3
umc1447	5.03	3	113-124	2	3	2	4
umc1545	7.00	4	69-85	3	2	3	3
umc1917	1.04	3	132-153	2	4	3	6
umc2047	1.09	4	100-134	7	7	7	8
umc2250	2.04	3	153-153	0	0	0	1

molecular analysis provides a wider genome sampling than the phenotypic analysis, therefore it is able to give a clear picture of genetic distance. The variation detected by the molecular markers is non-adaptive, hence not affected by natural or artificial selection. Most desirable phenotypic traits in plant breeding are a result of interaction among expressed genes, but agronomic studies are still essential in germplasm description and determination of molecular genetic distance is a complement (Donini et al., 2000). Clear estimates of the genetic distances would be closer when there is association between the loci controlling the phenotypic trait of interest (QTL) and the markers used and when a larger number of the traits of interest in relation to a particular situation are evaluated (Roy et al., 2004; Lefebvre et al., 2001). Earlier studies have reported that it is necessary to consider the molecular and phenotypic data separately in genotype divergence studies (Warburton et al., 2002). The use of phenotypic traits is therefore, relatively less efficient in discrimination of closely related genotypes and analysis of their genetic relationships compared to the use of molecular markers. Nevertheless, the use of phenotypic traits serves as a general approach in germplasm classification within a collection in relation to a particular trait.

The multivariate analyses revealed high concordance among the PCA, model-based population partitioning, clustering based on the genetic distance and discriminant analyses in terms of the number of groups and members in each group. Earlier studies have shown that principal component analysis as well as population structure are good predictors of grouping patterns and they can be used to complement the clustering method analysis, since different combinations of genetic distance matrices and clustering algorithms can give rise to somewhat different groups (Reif et al., 2005; Semagn et al., 2012).

The FST values form the analysis of molecular variance indicates a moderate genetic differentiation among groups and or populations. This is in agreement with the results of genetic diversity studies from previous research on maize populations (Semagn et al., 2012; Wen et al., 2012). In addition it has been reported that most variation in maize populations is partitioned within, rather than between populations, because maize is an out-crossing species a factor that lead to reduced population differentiation (Hamrick and Godt 1997).

Genetic divergence for resistance to stem borer and postharvest insect pests exists in tropical maize germplasm.

Appendix 2 Summary of group names of the clustering based on Rodgers genetic distance

Genotype and group name			
611D=SPRandotherlines	CML312=SPRandotherlines	CKPH08014=SPR-hybrids	CKIR09003=SBR-hybrids
CKSBL10003=SPRandotherlines	CML312-CML442=SPRandotherlines	CKPH08026=SPR-hybrids	CKIR06006=SBR-hybrids
CML488=SPRandotherlines	CKSBL10007=SPRandotherlines	CKPH08028=SPR-hybrids	CKIR09004=SBR-hybrids
SCDuma41=SPRandotherlines	CML511=SPRandotherlines	CKPH08024=SPR-hybrids	CKIR07010=SBR-hybrids
CKSBL10008=SPRandotherlines	CKSBL10004=SBR-lines	CKPH08003=SPR-hybrids	CKIR07013=SBR-hybrids
CKSBL10041=SPRandotherlines	CKSBL10026=SBR-lines	CKPH09004=SPR-hybrids	CKIR07005=SBR-hybrids
CKSBL10023=SPRandotherlines	CKSBL10025=SBR-lines	CKPH08004=SPR-hybrids	CKIR07001=SBR-hybrids
CKSPL10344=SPRandotherlines	CKSBL10005=SBR-lines	CKPH08002=SPR-hybrids	CKIR07004=SBR-hybrids
DTPWC9-49=SPRandotherlines	CKSBL10021=SBR-lines	CKPH08025=SPR-hybrids	CKIR07012=SBR-hybrids
CKSBL10035=SPRandotherlines	CKSBL10020=SBR-lines	CKPH08020=SPR-hybrids	CML264=SBR-hybrids
CKSBL10039=SPRandotherlines	CKSBL10027=SBR-lines	CKIR06008=SPR-hybrids	CKIR07018=SBR-hybrids
CKSBL10040=SPRandotherlines	CKSBL10001=SBR-lines	CKPH08039=SPR-hybrids	CKIR06004=SBR-hybrids
CKSPL10341=SPRandotherlines	CKSBL10082=SBR-lines	CKPH09002=SPR-hybrids	CKIR07011=SBR-hybrids
CKSPL10343=SPRandotherlines	CKSBL10028=SBR-lines	CKPH08009=SPR-hybrids	H6210=CH-hybrids
CKSPL10042=SPRandotherlines	CKSBL10030=SBR-lines	CKPH08010=SPR-hybrids	H6212=CH-hybrids
CKSPL10036=SPRandotherlines	CKSBL10029=SBR-lines	CKPH08012=SPR-hybrids	DK8031=CH-hybrids
CKSPL10021=SPRandotherlines	CKSBL10045=SBR-lines	CKPH08038=SPR-hybrids	H6213=CH-hybrids
CKSPL10035=SPRandotherlines	CML334=SBR-lines	CKPH09001=SPR-hybrids	H628=CH-hybrids
CKSPL10186=SPRandotherlines	CML442=SBR-lines	CKPH08040=SPR-hybrids	KH600-15A=CH-hybrids
CKSPL10224=SPRandotherlines	CKSBL10015=SBR-lines	CKPH08044=SPR-hybrids	H626=CH-hybrids
CKSPL10229=SPRandotherlines	CKSBL10013=SBR-lines	CKPH08033=SPR-hybrids	H629=CH-hybrids
CKSPL10295=SPRandotherlines	CKSBL10014=SBR-lines	CKPH08041=SPR-hybrids	DH01=CH-hybrids
CKSPL10309=SPRandotherlines	CKSBL10016=SBR-lines	CKPH08032=SPR-hybrids	H513=CH-hybrids
CKSPL10146=SPRandotherlines	CKSBL10060=SBR-lines	CKPH08036=SPR-hybrids	PH1=CH-hybrids
CKSPL10090=SPRandotherlines	CKSBL10004=SBR-lines	CKPH08037=SPR-hybrids	DH02=CH-hybrids
CKSPL10088=SPRandotherlines	CKSBL10026=SBR-lines	CKPH08043=SPR-hybrids	CKIR09007=CH-hybrids
CKSPL10089=SPRandotherlines	CKSBL10025=SBR-lines	CKPH08035=SPR-hybrids	DH04=CH-hybrids
CKSPL10164=SPRandotherlines	CKSBL10005=SBR-lines	CKPH09003=SPR-hybrids	
CKSPL10280=SPRandotherlines	CKSBL10021=SBR-lines	H6210=CH-hybrids	
CKSPL10256=SPRandotherlines	CKSBL10020=SBR-lines	H6212=CH-hybrids	
CKSPL10273=SPRandotherlines	CKSBL10027=SBR-lines	DK8031=CH-hybrids	
CKSPL10087=SPRandotherlines	CKSBL10001=SBR-lines	H6213=CH-hybrids	
CKSPL10136=SPRandotherlines	CKSBL10082=SBR-lines	H628=CH-hybrids	
CKSPL10230=SPRandotherlines	CKSBL10028=SBR-lines	KH600-15A=CH-hybrids	
CKSPL10086=SPRandotherlines	CKSBL10030=SBR-lines	H626=CH-hybrids	
CKSPL10074=SPRandotherlines	CKSBL10029=SBR-lines	H629=CH-hybrids	
CKSPL10080=SPRandotherlines	CKSBL10045=SBR-lines	DH01=CH-hybrids	
CKSPL10081=SPRandotherlines	CML334=SBR-lines	H513=CH-hybrids	
CKSPL10013=SPRandotherlines	CML442=SBR-lines	PH1=CH-hybrids	
CKSPL10212=SPRandotherlines	CKSBL10015=SBR-lines	DH02=CH-hybrids	
CKSPL10206=SPRandotherlines	CKSBL10013=SBR-lines	CKIR09007=CH-hybrids	
CKSPL10003=SPRandotherlines	CKSBL10014=SBR-lines	DH04=CH-hybrids	

Continuing Appendix 2

Genotype and group name		
CKSPL10218=SPRandotherlines	CKSBL10016=SBR-lines	CKIR04002=SBR-hybrids
CKSPL10113=SPRandotherlines	CKSBL10060=SBR-lines	CKIR04003=SBR-hybrids
CKSPL10111=SPRandotherlines	CKSBL10038=NA	CZL00003=SBR-hybrids
CKSPL10112=SPRandotherlines	CKSBL10033=NA	CKIR07003=SBR-hybrids
CKSPL10170=SPRandotherlines	CKSBL10042=NA	CML395=SBR-hybrids
CKSPL10177=SPRandotherlines	CKIR07009=NA	CKIR07008=SBR-hybrids
LaPosta-50=SPRandotherlines	PH3253=NA	CKIR07002=SBR-hybrids
LPSC7-52=SPRandotherlines	CML144=SPR-hybrids	CML395-CML444=SBR-hybrids
CML254=SPRandotherlines	500Q=SPR-hybrids	CKIR06001=SBR-hybrids
P100C-54=SPRandotherlines	631Q=SPR-hybrids	CKIR09002=SBR-hybrids
CML441=SPRandotherlines	SCDuma43=SPR-hybrids	CKIR06009=SBR-hybrids
CML443=SPRandotherlines	SCSimba61=SPR-hybrids	PH4=SBR-hybrids
CZL01005=SPRandotherlines	CML445=SPR-hybrids	CKIR09008=SBR-hybrids
CKSBL10046=SPRandotherlines	CKIR07017=SPR-hybrids	CML202-CML204=SBR-hybrids
CKSBL10043=SPRandotherlines	531A=SPR-hybrids	CKIR09001=SBR-hybrids
CML159=SPRandotherlines	533A=SPR-hybrids	CKIR06007=SBR-hybrids
CZL03014=SPRandotherlines	WH403=SPR-hybrids	CKIR09006=SBR-hybrids
LPSC7-51=SPRandotherlines	CML204=SPR-hybrids	CKIR09005=SBR-hybrids

Using the biophysical/bioassay traits which are adaptive, it was possible to discriminate the resistant from the susceptible but not according to their pedigree. The integrated analysis using SSR markers suggested that the maize germplasm was likely to be composed of four subpopulations (k = 3), one group of storage pest resistance lines, another group of stem borer resistance lines related to stem borer resistant hybrids, a third group of storage pest resistant hybrids and a fourth group constituting commercial hybrids from different seed companies within Kenya and a mixed group formed by the remaining genotypes. The grouping based on the SSR markers was highly consistent with the pedigree data. The results of this study can be directly used by breeding programs to better explore the genetic variability within the groups to develop new lines and between the groups to generate hybrids resistant to both field and postharvest insect pests in maize.

3 Materials and Methods

3.1 Evaluation for maize stem borer

A total of one hundred eighty four maize genotypes comprising of 100 inbred lines and 84 hybrids, from CIMMYT Kenya selected from CIMMYT Kenya breeding program was used in the study (Appendix 2). All the 184 genotypes and 36 checks (20 stem borer resistant and 16 susceptible checks) were evaluated for *Chilo partellus* and *Busseola fusca* resistance in October 2010 and April 2011 at the Kenyan agricultural research institute (KARI) stations both in Kiboko and Embu, Kenya. Kiboko is a dry and mid-altitude agro-ecological zone located at an elevation of 975 meters above sea level (masl), 37° 75 Ɛ and 2° 15´S. it has a sandy clay soil with an average annual rainfall of 530 mm and a mean minimum and maximum temperature of 14.3 and 35.1℃, respectively. Embu is a moist and mid-altitude zone located at an elevation of 1350 masl, and 37°42´ E and 0°49´S. Embu has a clay loam soil with an annual rainfall of 1,200 mm and a mean minimum and maximum temperature of 14.1 and 25℃, respectively.

Trials were planted in two-row plots of 5 m long at 0.25 m between hills and 0.75 m between rows using an alpha lattice design, with three replications per location. Two seeds were planted per hill and later thinned to one, giving a total plant density of 53,333 plants per hectare. In order to ensure a healthy crop, agronomic practices including weeding, fertilizer application and supplemental irrigation were done according to good agricultural practices. Each plot

Appendix 3 Summary of grouping of the genotypes based on Euclidean distance, population structure; stem borer and storage pest resistance

S/N	Germplasm	Entry	Name	Group based on Euclidean distance From PowerMarker	Group based on Structure at K=2	Group based on Structure at K=3	Group based on Structure at K=4	Group based on Structure at K=5
1	Inbred	5082	CKSBL10029	SBR lines	Pop1	Pop2	Pop1	Pop4
2	Inbred	5083	CKSBL10005	SBR lines	Pop1	Pop2	Pop1	Pop2
3	Inbred	5084	P300C5S1B	Unassigned	Pop1	Pop2	Pop1	Pop4
4	Inbred	5085	LPSC7-52	SPR and other lines	Pop1	Mixed	Pop1	Pop4
5	Inbred	5086	CKSPL10086	SPR and other lines	Pop2	Pop1	Pop3	Pop5
6	Inbred	5087	CKSBL10028	SBR lines	Pop1	Mixed	Pop1	Pop4
7	Inbred	5088	CKSBL10045	SBR lines	Pop1	Pop2	Pop1	Pop2
8	Inbred	5089	CKSBL10040	SPR and other lines	Pop1	Mixed	Pop1	Pop4
9	Inbred	5090	CKSBL10030	SBR lines	Pop1	Pop2	Pop1	Pop2
10	Inbred	5091	CML78	Unassigned	Pop1	Mixed	Pop1	Pop4
11	Inbred	5092	CZL03014	SPR and other lines	Pop1	Pop1	Pop1	Pop4
12	Inbred	5093	CKSBL10039	SPR and other lines	Pop1	Mixed	Pop1	Pop4
13	Inbred	5095	CKSPL10170	SPR and other lines	Pop2	Pop1	Pop3	Pop5
14	Inbred	5096	CKSBL10025	SBR lines	Pop1	Pop2	Pop1	Pop2
15	Inbred	5097	CML511	SPR and other lines	Pop1	Mixed	Pop1	Pop4
16	Inbred	5098	CML444	Unassigned	Pop1	Mixed	Mixed	Mixed
17	Inbred	5099	CKSPL10036	SPR and other lines	Pop2	Pop1	Pop3	Pop5
18	Inbred	5100	CKSPL10088	SPR and other lines	Pop2	Pop1	Pop3	Pop5
19	Inbred	5101	CKSPL10111	SPR and other lines	Pop2	Pop1	Pop3	Pop5
20	Inbred	5102	CKSPL10113	SPR and other lines	Pop2	Pop1	Pop3	Pop5
21	Inbred	5103	CML440	Unassigned	Pop1	Mixed	Pop1	Pop4
22	Inbred	5104	CML334	SBR lines	Pop1	Mixed	Pop1	Pop4
23	Inbred	5105	CKSBL10042	Unassigned	Pop1	Pop2	Pop1	Mixed
24	Inbred	5106	CML442	SBR lines	Pop1	Mixed	Pop1	Pop4
25	Inbred	5107	P100C-54	SPR and other lines	Pop1	Mixed	Pop1	Pop4
26	Inbred	5108	CKSBL10033	Unassigned	Pop1	Pop2	Mixed	Mixed
27	Inbred	5109	CKSPL10218	SPR and other lines	Pop2	Pop1	Pop3	Pop5
28	Inbred	5110	CKSPL10089	SPR and other lines	Pop2	Pop1	Pop3	Pop5
29	Inbred	5111	CKSPL10090	SPR and other lines	Pop2	Pop1	Pop3	Pop5
30	Inbred	5112	CML443	SPR and other lines	Pop1	Mixed	Pop1	Pop4
31	Inbred	5114	CML144	Unassigned	Pop1	Mixed	Pop1	Pop4
32	Inbred	5115	CKSBL10027	SBR lines	Pop1	Mixed	Pop1	Pop2
33	Inbred	5116	CKSBL10046	SPR and other lines	Pop1	Mixed	Pop1	Pop4
34	Inbred	5117	LaPosta-50	SPR and other lines	Pop1	Mixed	Pop1	Pop4
35	Inbred	5118	CKSBL10007	SPR and other lines	Pop1	Mixed	Pop1	Pop4
36	Inbred	5119	CKSPL10177	SPR and other lines	Pop2	Pop1	Pop3	Pop5
37	Inbred	5120	CKSPL10344	SPR and other lines	Pop1	Pop2	Pop1	Pop4
38	Inbred	5121	DTPWC9-49	SPR and other lines	Pop1	Pop1	Pop1	Pop4
39	Inbred	5123	DTPWC9-48	Unassigned	Pop1	Mixed	Pop1	Pop4
40	Inbred	5124	CZL00003	Unassigned	Pop1	Pop2	Mixed	Mixed

Continuing Appendix 3

S/N	Germplasm	Entry	Name	Group based on Euclidean distance From PowerMarker	Group based on Structure at K=2	Group based on Structure at K=3	Group based on Structure at K=4	Group based on Structure at K=5
41	Inbred	5125	CKSPL10035	SPR and other lines	Pop2	Pop1	Pop3	Pop5
42	Inbred	5126	CKSBL10082	SBR lines	Pop1	Pop2	Mixed	Pop2
43	Inbred	5127	CKSBL10001	SBR lines	Pop1	Pop2	Pop1	Pop2
44	Inbred	5128	CKSBL10015	SBR lines	Pop1	Pop2	Pop1	Mixed
45	Inbred	5129	CML441	SPR and other lines	Pop1	Mixed	Pop1	Pop4
46	Inbred	5130	CKSBL10004	SBR lines	Pop1	Pop2	Pop1	Pop2
47	Inbred	5132	CKSPL10280	SPR and other lines	Pop2	Pop1	Pop3	Pop5
48	Inbred	5133	CKSBL10035	SPR and other lines	Pop1	Mixed	Pop1	Pop4
49	Inbred	5134	CKSBL10014	SBR lines	Pop1	Mixed	Pop1	Pop4
50	Inbred	5135	CKSPL10273	SPR and other lines	Pop2	Pop1	Pop3	Pop5
51	Inbred	5136	CKSBL10013	SBR lines	Pop1	Mixed	Pop1	Pop4
52	Inbred	5137	CKSPL10042	SPR and other lines	Pop2	Pop1	Pop3	Pop5
53	Inbred	5138	CKSPL10256	SPR and other lines	Pop2	Pop1	Pop3	Pop5
54	Inbred	5139	CKSPL10003	SPR and other lines	Pop2	Pop1	Pop3	Pop5
55	Inbred	5140	CML488	SPR and other lines	Pop1	Mixed	Pop1	Pop4
56	Inbred	5141	CKSPL10230	SPR and other lines	Pop2	Pop1	Pop3	Pop5
57	Inbred	5142	CKSBL10021	SBR lines	Pop1	Pop2	Pop1	Pop2
58	Inbred	5143	CKSPL10309	SPR and other lines	Pop2	Pop1	Pop3	Pop5
59	Inbred	5144	CML264	Unassigned	Pop1	Pop2	Mixed	Mixed
60	Inbred	5146	CKSBL10038	Unassigned	Pop1	Pop2	Pop1	Pop4
61	Inbred	5147	CKSBL10020	SBR lines	Pop1	Mixed	Pop1	Pop2
62	Inbred	5148	CKSBL10043	SPR and other lines	Pop1	Mixed	Pop1	Pop4
63	Inbred	5149	CKSPL10021	SPR and other lines	Pop2	Pop1	Pop3	Pop5
64	Inbred	5150	CML159	SPR and other lines	Pop1	Mixed	Pop1	Pop4
65	Inbred	5151	CKSPL10164	SPR and other lines	Pop2	Pop1	Pop3	Pop5
66	Inbred	5153	CKSPL10343	SPR and other lines	Pop1	Mixed	Pop1	Pop4
67	Inbred	5155	CKSPL10080	SPR and other lines	Pop2	Pop1	Pop3	Pop5
68	Inbred	5157	CKSPL10212	SPR and other lines	Pop2	Pop1	Pop3	Pop5
69	Inbred	5158	CKSBL10026	SBR lines	Pop1	Pop2	Pop1	Pop2
70	Inbred	5160	DTPWC9-53	Unassigned	Pop1	Mixed	Pop1	Pop4
71	Inbred	5161	CKSBL10041	SPR and other lines	Pop1	Pop1	Pop1	Pop4
72	Inbred	5162	CML204	Unassigned	Pop2	Mixed	Mixed	Mixed
73	Inbred	5163	CML445	Unassigned	Pop1	Mixed	Pop1	Pop4
74	Inbred	5164	CKSPL10206	SPR and other lines	Pop2	Pop1	Pop3	Pop5
75	Inbred	5165	CKSBL10060	SBR lines	Pop1	Mixed	Pop1	Pop4
76	Inbred	5166	CKSBL10023	SPR and other lines	Pop1	Mixed	Pop1	Pop4
77	Inbred	5167	CKSPL10224	SPR and other lines	Pop2	Pop1	Pop3	Pop5
78	Inbred	5168	CKSPL10087	SPR and other lines	Pop2	Pop1	Pop3	Pop5
79	Inbred	5171	CKSPL10341	SPR and other lines	Pop1	Pop1	Mixed	Pop4
80	Inbred	5172	CKSPL10295	SPR and other lines	Pop2	Pop1	Pop3	Pop5
81	Inbred	5173	CKSPL10112	SPR and other lines	Pop2	Pop1	Pop3	Pop5

Continuing Appendix 3

S/N	Germplasm	Entry	Name	Group based on Euclidean distance From PowerMarker	Group based on Structure at K=2	Group based on Structure at K=3	Group based on Structure at K=4	Group based on Structure at K=5
82	Inbred	5174	CKSPL10186	SPR and other lines	Pop2	Pop1	Pop3	Pop5
83	Inbred	5175	CKSPL10146	SPR and other lines	Pop2	Pop1	Pop3	Pop5
84	Inbred	5176	CZL03007	Unassigned	Mixed	Mixed	Pop1	Pop4
85	Inbred	5179	CZL01005	SPR and other lines	Pop1	Pop2	Pop1	Pop4
86	Inbred	5180	CKSBL10003	SPR and other lines	Pop1	Pop2	Mixed	Mixed
87	Inbred	5182	CKSPL10013	SPR and other lines	Pop2	Pop1	Pop3	Pop5
88	Inbred	5183	CKSPL10081	SPR and other lines	Pop2	Pop1	Pop3	Pop5
89	Inbred	5184	CKSBL10016	SBR lines	Pop1	Mixed	Pop1	Pop4
90	Inbred	5185	CKSPL10136	SPR and other lines	Pop2	Pop1	Pop3	Pop5
91	Inbred	5186	CML489	Unassigned	Pop1	Mixed	Pop1	Pop4
92	Inbred	5187	LPSC7-51	SPR and other lines	Pop1	Mixed	Pop1	Pop4
93	Inbred	5188	CML395	Unassigned	Pop1	Mixed	Pop1	Mixed
94	Inbred	5189	CML254	SPR and other lines	Mixed	Pop1	Pop1	Pop4
95	Inbred	5190	CML202	Unassigned	Mixed	Mixed	Mixed	Mixed
96	Inbred	5191	CKSPL10074	SPR and other lines	Pop2	Pop1	Pop3	Pop5
97	Inbred	5192	CML312	SPR and other lines	Pop1	Mixed	Pop1	Pop4
98	Inbred	5193	CML197	Unassigned	Pop1	Pop1	Pop1	Pop4
99	Inbred	5194	CKSBL10008	SPR and other lines	Pop1	Pop1	Pop1	Pop4
100	Inbred	5195	CKL06-1	Unassigned	Pop1	Mixed	Pop1	Pop4
101	Inbred	5196	CKSPL10229	SPR and other lines	Pop2	Pop1	Pop3	Pop5
102	Hybrid	5001	CKIR09005	SBR hybrids	Pop1	Pop2	Pop2	Pop1
103	Hybrid	5002	CKPH08012	SPR hybrids	Pop2	Pop3	Pop4	Pop3
104	Hybrid	5003	CKIR07013	SBR hybrids	Pop1	Pop2	Mixed	Mixed
105	Hybrid	5004	CKIR07009	Unassigned	Pop1	Pop2	Pop2	Pop1
106	Hybrid	5005	PH3253	Unassigned	Pop1	Pop2	Mixed	Pop4
107	Hybrid	5006	CKPH08038	SPR hybrids	Pop2	Pop3	Pop4	Pop3
108	Hybrid	5007	CML395-CML444	SBR hybrids	Pop1	Pop2	Pop2	Pop1
109	Hybrid	5008	CML312-CML442	SPR and other lines	Pop1	Pop2	Pop1	Pop4
110	Hybrid	5009	CKPH09002	SPR hybrids	Pop2	Pop3	Pop4	Pop3
111	Hybrid	5010	SCDuma43	SPR hybrids	Pop1	Pop2	Mixed	Mixed
112	Hybrid	5011	DH01	CH susceptible	Pop1	Pop2	Pop2	Pop1
113	Hybrid	5012	CKPH09004	SPR hybrids	Pop2	Pop3	Pop4	Pop3
114	Hybrid	5013	CKIR09006	SBR hybrids	Pop1	Pop2	Pop2	Pop1
115	Hybrid	5014	H6210	CH susceptible	Pop1	Pop2	Pop2	Pop1
116	Hybrid	5015	H628	CH susceptible	Pop1	Pop2	Pop2	Pop1
117	Hybrid	5016	H629	CH susceptible	Pop1	Pop2	Pop2	Pop1
118	Hybrid	5017	CKPH08010	SPR hybrids	Pop2	Pop3	Pop4	Pop3
119	Hybrid	5018	CKIR09008	SBR hybrids	Pop1	Pop2	Pop2	Pop1
120	Hybrid	5019	CKPH08004	SPR hybrids	Pop2	Pop3	Pop4	Pop3
121	Hybrid	5020	CKPH08037	SPR hybrids	Pop2	Pop3	Pop4	Pop3

Continuing Appendix 3

S/N	Germplasm	Entry	Name	Group based on Euclidean distance From PowerMarker	Group based on Structure at K=2	Group based on Structure at K=3	Group based on Structure at K=4	Group based on Structure at K=5
122	Hybrid	5021	CKPH08043	SPR hybrids	Pop2	Pop3	Pop4	Pop3
123	Hybrid	5022	CKIR07008	SBR hybrids	Pop1	Pop2	Pop2	Pop1
124	Hybrid	5023	CKIR07005	SBR hybrids	Pop1	Pop2	Pop2	Pop1
125	Hybrid	5024	CKIR09002	SBR hybrids	Pop1	Pop2	Pop2	Pop1
126	Hybrid	5025	CKPH08024	SPR hybrids	Pop2	Pop3	Pop4	Pop3
127	Hybrid	5026	CKPH08036	SPR hybrids	Pop2	Pop3	Pop4	Pop3
128	Hybrid	5027	CKPH08032	SPR hybrids	Pop2	Pop3	Pop4	Pop3
129	Hybrid	5028	CKIR07003	SBR hybrids	Pop1	Pop2	Mixed	Mixed
130	Hybrid	5029	CKIR07010	SBR hybrids	Pop1	Pop2	Pop2	Pop1
131	Hybrid	5030	CKIR07018	SBR hybrids	Pop1	Pop2	Pop2	Pop1
132	Hybrid	5031	CKIR07011	SBR hybrids	Pop1	Pop2	Pop2	Pop1
133	Hybrid	5032	CKPH08033	SPR hybrids	Pop2	Pop3	Pop4	Pop3
134	Hybrid	5033	CKIR06009	SBR hybrids	Pop1	Pop2	Pop2	Pop1
135	Hybrid	5034	CKPH08035	SPR hybrids	Pop2	Pop3	Pop4	Pop3
136	Hybrid	5035	631Q	SPR hybrids	Pop1	Mixed	Mixed	Mixed
137	Hybrid	5036	CKPH08028	SPR hybrids	Pop2	Pop3	Pop4	Pop3
138	Hybrid	5037	WH403	SPR hybrids	Pop1	Mixed	Mixed	Mixed
139	Hybrid	5038	H6212	CH susceptible	Pop1	Pop2	Pop2	Pop1
140	Hybrid	5039	CKIR06006	SBR hybrids	Pop1	Pop2	Pop2	Pop1
141	Hybrid	5040	CKPH08009	SPR hybrids	Pop2	Pop3	Pop4	Pop3
142	Hybrid	5041	CKPH08014	SPR hybrids	Pop2	Pop3	Pop4	Pop3
143	Hybrid	5042	CKIR09003	SBR hybrids	Pop1	Pop2	Pop2	Pop1
144	Hybrid	5043	533A	SPR hybrids	Pop1	Mixed	Mixed	Mixed
145	Hybrid	5044	CKPH09001	SPR hybrids	Pop2	Pop3	Pop4	Pop3
146	Hybrid	5045	H6213	CH susceptible	Pop1	Pop2	Pop2	Pop1
147	Hybrid	5046	531A	SPR hybrids	Pop1	Mixed	Mixed	Mixed
148	Hybrid	5047	CML202-CML204	SBR hybrids	Pop1	Pop2	Pop2	Pop1
149	Hybrid	5048	CKIR06008	SPR hybrids	Pop2	Pop3	Pop4	Pop3
150	Hybrid	5049	CKIR07001	SBR hybrids	Pop1	Pop2	Pop2	Pop1
151	Hybrid	5050	CKIR06001	SBR hybrids	Pop1	Pop2	Pop2	Pop1
152	Hybrid	5051	CKIR07012	SBR hybrids	Pop1	Pop2	Mixed	Mixed
153	Hybrid	5052	CKPH09003	SPR hybrids	Pop2	Pop3	Pop4	Pop3
154	Hybrid	5053	CKPH08044	SPR hybrids	Pop2	Pop3	Pop4	Pop3
155	Hybrid	5054	CKPH08020	SPR hybrids	Pop2	Pop3	Pop4	Pop3
156	Hybrid	5055	611D	Unassigned	Pop1	Pop2	Mixed	Mixed
157	Hybrid	5056	DK8031	CH susceptible	Pop1	Pop2	Pop2	Pop1
158	Hybrid	5057	CKPH08025	SPR hybrids	Pop2	Pop3	Pop4	Pop3
159	Hybrid	5058	CKPH08040	SPR hybrids	Pop2	Pop3	Pop4	Pop3
160	Hybrid	5059	CKPH08039	SPR hybrids	Pop2	Pop3	Pop4	Pop3
161	Hybrid	5060	SCSimba61	SPR hybrids	Pop1	Pop2	Mixed	Mixed
162	Hybrid	5061	CKPH08026	SPR hybrids	Pop2	Pop3	Pop4	Pop3

Continuing Appendix 3

S/N	Germplasm	Entry	Name	Group based on Euclidean distance From PowerMarker	Group based on Structure at K=2	Group based on Structure at K=3	Group based on Structure at K=4	Group based on Structure at K=5
163	Hybrid	5062	CKIR09001	SBR hybrids	Pop1	Pop2	Pop2	Pop1
164	Hybrid	5063	PH4	SBR hybrids	Pop1	Pop2	Pop2	Pop1
165	Hybrid	5064	CKIR09007	CH susceptible	Pop1	Pop2	Pop2	Pop1
166	Hybrid	5065	H626	CH susceptible	Pop1	Pop2	Pop2	Pop1
167	Hybrid	5066	PH1	CH susceptible	Pop1	Pop2	Pop2	Pop1
168	Hybrid	5067	CKPH08003	SPR hybrids	Pop2	Pop3	Pop4	Pop3
169	Hybrid	5068	H513	CH susceptible	Pop1	Pop2	Pop2	Pop1
170	Hybrid	5069	CKIR09004	SBR hybrids	Pop1	Pop2	Pop2	Pop1
171	Hybrid	5070	CKIR07004	SBR hybrids	Pop1	Pop2	Pop2	Mixed
172	Hybrid	5071	CKIR07017	SPR hybrids	Pop1	Mixed	Pop2	Pop1
173	Hybrid	5072	CKIR06004	SBR hybrids	Pop1	Pop2	Pop2	Pop1
174	Hybrid	5073	KH600-15A	CH susceptible	Pop1	Pop2	Pop2	Pop1
175	Hybrid	5074	SCDuma41	SPR and other lines	Pop1	Mixed	Pop1	Pop4
176	Hybrid	5075	500Q	SPR hybrids	Pop1	Mixed	Mixed	Mixed
177	Hybrid	5076	CKPH08002	SPR hybrids	Pop2	Pop3	Pop4	Pop3
178	Hybrid	5077	CKIR06007	SBR hybrids	Pop1	Pop2	Pop2	Pop1
179	Hybrid	5078	CKPH08041	SPR hybrids	Pop2	Pop3	Pop4	Pop3
180	Hybrid	5079	DH04	CH susceptible	Pop1	Pop2	Pop2	Pop1
181	Hybrid	5080	DH02	CH susceptible	Pop1	Pop2	Pop2	Pop1
182	Hybrid	5081	CKIR07002	SBR hybrids	Pop1	Pop2	Pop2	Pop1
183	Hybrid	5274	CKIR04003	SBR hybrids	Pop1	Pop2	Mixed	Mixed
184	Hybrid	5280	CKIR04002	SBR hybrids	Pop1	Pop2	Pop2	Pop1

was divided into two equal parts, one section for 10 stem borer infested plants while the other portion of 10 plants was protected from stem borer damage by applying bulldock® 25 EC insecticide at a concentration of 25 g/l Beta-Cyfluthrin and acted as the control. Stem borer infestation was done approximately 3 weeks after planting by artificially infesting the 10 plants per plot with 5 first-instar neonates of *C. partellus* in Kiboko and *B. fusca* in Embu. The stem borers larvae used in this experiment were obtained from KARI-Katumani insectary. Leaf-damage for each individual plant was scored two weeks after infestation on a scale of 1 to 9 (1= no visible leaf damage; 9= plants dying as a result of leaf damage) (Tefera et al., 2011). At harvest, the numbers of exit holes on the stems were counted and the cumulative tunnel length was measured by splitting the stems. Ears from stem borer uninfested plots were harvested, sun-dried to a moisture content of 12-13 % and used for, maize weevil and larger grain borer evaluation at the KARI/CIMMYT Entomology Laboratory in Kiboko as described below.

3.2 Evaluation for maize weevil and larger grain borer

The maize grains were disinfested by fumigating with phostoxin tablets for seven days to eliminate field infestation. For each genotype 100 grams of grain from each plot per replication was placed in 250 ml jars, infested with 50 unsexed 7-10 day old maize weevils and larger grain borer separately, and stored for 90 days at a temperature of 26-28 °C and relative humidity of 70-75 %. The insects used in the experiment were obtained from the KARI/CIMMYT Kiboko maize Entomology Laboratory where they were reared on the grains of maize cultivar PH3253 under controlled conditions (28 °C and 75% relative humidity). Evaluation was conducted using a completely randomized design with 3 replications.

The contents of each jar were sieved with mesh (Endecotts Ltd, UK[1]) 90 days after infestation to separate grains, insects and flour. The flour produced by the insects was weighed, while the number of damaged kernels and adult insect progeny were counted. The grain weight loss was computed by subtracting the final from the initial weight of the grain sample and expressed as a percentage (Tefera et al. 2011). Damaged kernels were separated from the undamaged based on grain tunnelling and holes. The percentage of damaged grain was computed. Finally, the weight of the damaged and undamaged grains was measured.

3.3 DNA extraction and genotyping

Leaf samples were harvested from 10 healthy plants per genotype about 3 weeks after sowing at the Kiboko station. They were sampled in perforated Ziploc bags, immediately transferred into a Styrofoam box containing dry ice and transported to the Biosciences for eastern and central Africa (BecA) hub in Nairobi. Approximately equal amount of leaf tissue from each of the 10 plants per genotype was bulked, cut into pieces, and transferred into 1.2 ml strip tubes that contained two 4-mm stainless steel grinding balls (Spex CetriPrep, USA). The leaf samples were freeze-dried for 4 days using a Labconco freeze dryer (http://www.labconco.com) as described in the user's manual. The lyophilized leaf samples were ground into fine powder at 1500 strokes per minute for 2 minutes using GenoGrinder-2000 and genomic DNA was extracted using a modified version of the CIMMYT high throughput mini-prep Cetyl Trimethyl Ammonium Bromide (CTAB) method as described elsewhere (Semagn 2014). The quality of the isolated DNA was checked after running aliquots of DNA samples on a 0.8% agarose gel that contained 0.3 µg/mL Gel-Red-(Biotium). DNA concentration was measured using NanoDrop-ND-1000 Spectrophotometer, (Thermo Scientific, Wilmington, DE 19810, USA).

The samples were genotyped with 56 fluorescently-labelled SSRs (Appendix 1), selected from the list of markers used for the genetic characterization of CIMMYT maize inbred lines and OPVs (Warburton et al., 2002). Polymerase Chain Reaction (PCR), genotyping and data scoring were done as described in another paper (Semagn et al., 2014). Both DNA extraction and genotyping were done at the Biosciences Eastern and Central Africa (BecA) hub.

3.4 Analysis of phenotypic data

The percentage weight loss, flour weight and grain damage data were transformed using arcsine transformation to normalize its frequency distribution. A univariate analysis of variance using the general linear model (GLM) procedure of SAS version 9.3 (SAS Institute 2003) was performed on grain biophysical and insect bioassay traits as well as the stem borer damage traits. A susceptibility index based on leaf damage score, number of borer exit holes and cumulative tunnel length was computed by summing up the ratios between values and overall mean and dividing by the number of parameters evaluated. Germplasm with susceptibility-index values less than 0.8 were regarded as resistant, and those with greater than 0.8 as susceptible (Tefera et al. 2011).

3.5 Analysis of molecular data

SSR data analyses were conducted as described by Semagn et al., (2014). Briefly, AlleloBin *(http://www.icrisat.org/bt-software-downloads.htm*) was used for adjusting inconsistencies in allele calls obtained from GeneMapper software. The number of adjusted alleles per locus for each bulked genotype varied from 2 to 11. Thus, the adjusted allele sizes were converted into binary format (present =1 and absent = 0) using ALS-Binary*(http://www.icrisat.org/bt-software-downloads.htm*). Rogers distance matrix was calculated between each pair of genotypes using NTSYS-pc for Windows, version 2.0. The distance matrix was used to generate phenograms using the unweighted pair-group method based on arithmetic average (UPGMA) as implemented in MEGA5.1. Principal component analysis (PCA) was performed to project the genotypes into different groups using JMP version 7.0 (SAS institute Inc., Cary, NC, USA). The first two principal components were plotted to visualize patterns of relationships among genotypes. An admixture model-based clustering method implemented in the software package STRUCTURE version 2.3.3 (Pritchard et al., 2000) was used to infer population structure among genotypes. STRUCTURE was run by varying the number of clusters (k) from 1 to 6, with each K repeated thrice at a burn-in period of 100,000 and 100,000 MCMC (Markov Chain Monte Carlo) replications after burn-in. Genotypes with membership probabilities > 60% were assigned to the same group,

while those with < 60% probability memberships in any single groups were assigned to a "mixed" group. A stepwise forward canonical discriminant analysis was run using SAS statistical package (SAS Institute 2003). Analysis of molecular variance (AMOVA) was used to partition the variation among and within groups using ARLEQUIN version 3.11. For both discriminant analysis and AMOVA, the genotypes were assigned into groups or populations using the results from the phenotypic data, STRUCTURE and cluster analysis (Appendix 3).

Authors' contributions

All the authors participated in carrying out the experiments towards generation of data. S assisted in handling of the molecular data. JK, All the co-authors were involved in drafting the manuscript and reviewing it for quality check. K, S and P were the key supervisors of the research work as part of advisory team for PhD research work.

Acknowledgement

The financial support by the Regional Universities Forum for Capacity building in Agriculture (RUFORUM), International Maize and Wheat Improvement Center (CIMMYT) and the Syngenta Foundation for Sustainable Agriculture through the Insect Resistant Maize for Africa (IRMA) Project is gratefully acknowledged. Mr. Joel Mbithi, Andrew Chavangi, Mr. Patrick Gichobi, Mr. Charles Marangu, Mrs.Veronica Ogungo, the technical staff at KARI-Kiboko; and Mr. David Karuri, of KARI-Embu, are recognized for their efforts in assisting with the field trials and laboratory experiments.

References

Arnason J.T., Baum J., Gale J.D., Lambert R., Bergvinson D., Philogene B.J., Serratos J.A., Mihm J., and Jewell D.C., 1994, Variation in resistance of Mexican landraces of maize to maize weevil *Sitophilus zeamais*, in relation to taxonomic and biochemical parameters, Euphytica, 74: 227-236
http://dx.doi.org/10.1007/BF00040405

Bertan I., Carvalho F., and Oliveria A., 2007, Parental selection strategies in plant breeding programs, Journal of Crop Science and Biotechnology, 10: 211-222

Bohn M., Utz H., and Melchinger A., 1999, Genetic similarities among winter wheat cultivars determined on the basis of RFLPs, AFLPs, SSRs and their use for predicting progeny variance, Crop Science, 39: 228-237
http://dx.doi.org/10.2135/cropsci1999.0011183X003900010035x

Dhliwayo T., and Pixley K.V., 2003, Divergent selection for resistance to maize weevil in six maize populations, Crop Science, 43: 2043-2049
http://dx.doi.org/10.2135/cropsci2003.2043

Donini P., Law J., Koebner R., and Reeves J., 2000, Temporal trends in the diversity of UK wheat, Theoretical Applied Genetics, 100: 912-917
http://dx.doi.org/10.1007/s001220051370

Gupta P., Balyan H., Edwards K., Isaac P., Korzun V., Röder M., Gautier M.F., Joudrier P., Schlatter A., Dubcovsky J., De la Pena R., Khairallah M., Penner G., Hayden M., Sharp P., Keller B., Wang R., Hardouin J., Jack P., and Leroy P., 2002, Genetic mapping of 66 new microsatellite (SSR) loci in bread wheat, Theoretical and Applied Genetics, 105: 413-422
http://dx.doi.org/10.1007/s00122-002-0865-9

Hamblin M., Warburton M., and Buuckle E., 2007, Empirical comparison of simple sequence repeats and single nucleotide polymorphisms in assessment of maize diversity and relatedness, PloS One, 2: e1367
http://dx.doi.org/10.1371/journal.pone.0001367

Hamrick J.L., and Godt M.J.W., 1997, Allozyme diversity in cultivated crops, Crop Science, 37: 26-30
http://dx.doi.org/10.2135/cropsci1997.0011183X003700010004x

Koebner R., Donini P., Reeves J., Cooke, R., and Law, J., 2002. Temporal flux in the morphological and molecular diversity of UK barley, Theoretical Applied Genetics, 106, 550-558

Lefebvre V., Goffinet B., Chauvet J., Caromel B., Signoret P., and Brand R., 2000, Evaluation of genetic distances between pepper inbred lines for cultivar protection purposes: Comparison of AFLP, RAPD, and phenotypic data, Theoretical Applied Genetics, 102: 741-750
http://dx.doi.org/10.1007/s001220051705

Maric S., Bolaric S., Martinc˘ic J., Pejic´ I., and Kozumplik V., 2004, Genetic diversity of hexaploid wheat cultivars estimated by RAPD markers, morphological traits and coefficients of parentage, Plant Breeding, 123: 366-369
http://dx.doi.org/10.1111/j.1439-0523.2004.00956.x

Mohammadi S., and Prasanna B., 2003, Analysis of genetic diversity in crop plants-salient statistical tools and consideration, Crop Science, 43: 1235-1248
http://dx.doi.org/10.2135/cropsci2003.1235

Mukhtar M., Rahman M., and Zafar Y., 2002, Assessment of genetic diversity among wheat (*Triticum aestivum* L.) cultivars from a range of localities across Pakistan using random amplified polymorphic DNA (RAPD) analysis, Euphytica, 128: 417-425
http://dx.doi.org/10.1023/A:1021261811454

Munyiri S.W., Okori P., Mugo S.N., Otim M., Gibson P., and Mwololo J.K., 2010, Genetic diversity in maize landraces for resistance to *Chilo partellus* in Kenya, In: Second RUFORUM Biennial Conference proceedings held on 20- 24 September, 2010, Entebbe, Uganda, pp.1003-1007

Mwololo J.K., Mugo S., Okori P., Tefera T., and Munyiri S.W., 2010, Genetic diversity for resistance to larger grain borer in maize hybrids and open pollinated varieties in Kenya, Second RUFORUM Biennial Meeting, Kampala, Uganda, pp.535-539

Prasanna B., Pixley K., Warburton M., and Xie C.X., 2010, Molecular marker-assisted breeding options for maize improvement in Asia, Molecular Breeding, 26: 339-356
http://dx.doi.org/10.1007/s11032-009-9387-3

Pritchard J.K., Stephens M., and Donnelly P., 2000, Inference of population structure using multilocus genotype data, Genetics, 155: 945-959

Reif J.C., Melchinger A.E., and Frish M., 2005, Genetical and mathematical properties of similarity and dissimilarity coefficient applied in plant breeding and seed bank management, Crop Sci., 45: 1-7
http://dx.doi.org/10.2135/cropsci2005.0001

Roy J., Lakshmikumaran M., Balyan H., and Gupta P., 2004, AFLP-based genetic diversity and its comparison with diversity based on SSR, SAMPL, and phenotypic traits in bread wheat, Biochemical Genetics, 42: 43-59
http://dx.doi.org/10.1023/B:BIGI.0000012143.48298.71

SAS Institute, 2003, Statistical software system, Proceedings of the Twenty-Eighth Annual SAS® Users Group International Conference, Cary, NC, USA

Semagn K., 2014, Leaf tissue sampling and DNA extraction protocols. In: Besse P.(ed.), Molecular plant taxonomy: methods and protocols, Human Press, New York, pp.53-67
http://dx.doi.org/10.1007/978-1-62703-767-9_3

Semagn K., Magorokosho C., Ogugo V., Makumbi D., and Warburton M.L., 2014, Genetic relationships and structure among open-pollinated maize varieties adapted to eastern and southern Africa using microsatellite markers, Molecular Breeding, pp.1-13
http://dx.doi.org/10.1007/s11032-013-9917-x

Semagn K., Magorokosho C., Vivek B.S., Makumbi D., Beyene Y., Mugo S., Prasanna B.M., and Warburton M.L., 2012, Molecular characterization

of diverse CIMMYT maize inbred lines from eastern and southern Africa using single nucleotide polymorphic markers., BMC Genomics, 13:113. doi:10. 1186/1471-2164-13-113

Shiferaw B., Prasanna BM., Hellin J., and Bänziger M., 2011, Crops that feed the world, past successes and future challenges to the role played by maize in global food security, Food Security, 3: 307-327 http://dx.doi.org/10.1007/s12571-011-0140-5

Smith J., and Smith O., 1989, Comparison of heterosis among hybrids as a measure of relatedness with that to be expected on the basis of pedigree, Maize Genetics Cooperation News, 63: 86-87

Tefera T., 2012, Post-harvest losses in African maize in the face of increasing food shortage, Food security, 4: 267-277

Tefera T., Mugo S., Tende R., and Likhayo P., 2011, Methods of screening maize for resistance to stem borers and post-harvest insect pets, ISBN: 978-92-9059-289-1 ed, CIMMYT, Nairobi, Kenya

Warburton M., Skovmand B., and Mujeeb-Kazi, 2002, The molecular genetic characterization of the 'Bobwhite' bread wheat family using AFLPs and the effect of the T1BL.1RS translocation, Theoretical Applied Genetics 104: 868-873 http://dx.doi.org/10.1007/s00122-001-0816-x

Wen Y., John N., Eleazar E., and Eran H., 2012, A model-based approach for analysis of spatial structure in genetic data, Nature Genetics, 44: 725-732 http://dx.doi.org/10.1038/ng.2285

Xia X.C., Reif J.C., Melchinger A.E., Frisch M., Hoisington D.A., Beck D., Pixley K., and Warburton M.L., 2005, Genetic diversity among CIMMYT maize inbred lines investigated with SSR markers: II. Subtropical, tropical mid-latitude and highland maize inbred lines and their relationships with elite U.S. and European maize, Crop Science, 45: 2573-2582 http://dx.doi.org/10.2135/cropsci2005.0246

Gupta, Balyan, Edwards, Isaac, Korzun, Roder, Gautier, Joudrier, Schlatter, Dubcovsky, De L.P., Khairallah, Penner, Hayden, Sharp, Keller, Wang, Hardouin, Jack, and Leroy, 2002, Genetic mapping of 66 new microsatellite (SSR) loci in bread wheat, Theoretical and Applied Genetics, 105: 413-422

Prasanna B., Pixley K., Warburton M., and Xie C.X., 2010, Molecular marker-assisted breeding options for maize improvement in Asia, Molecular Breeding, 26: 339-356